DOCUMENT RESUME

ED 356 456 CS 011 268

AUTHOR Rosenshine, Barak; Meister, Carla E.
TITLE Reciprocal Teaching: A Review of 19 Experimental Studies. Technical Report No. 574.
INSTITUTION Center for the Study of Reading, Urbana, IL.
PUB DATE May 93
NOTE 54p.
PUB TYPE Reports - Research/Technical (143) -- Information Analyses (070)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Elementary Secondary Education; *Instructional Effectiveness; Literature Reviews; *Reading Comprehension; Reading Instruction; *Reading Research; *Reciprocal Teaching; Standardized Tests; Teacher Behavior; *Teaching Methods
IDENTIFIERS Alternative Assessment

ABSTRACT

Since the introduction of the concept of reciprocal teaching by A. S. Palincsar and A. L. Brown in 1984, researchers have conducted a number of studies that attempted to replicate and/or extend the original one. A review was conducted of 19 studies that used reciprocal teaching to help students improve their ability to comprehend text. The review found that: (1) results were usually more significant when explicit instruction in the cognitive strategies was provided before the reciprocal teaching began than when reciprocal teaching only was used; (2) results were mostly non-significant when below-average students were taught, yet usually significant when all other students were taught; and (3) results were usually significant when experimenter-developed tests were used, yet usually non-significant when standardized tests were used. (Fourteen tables of data are included.) Two appendixes provide lists of studies analyzed and a script illustrating the explicit teaching of questioning prior to the reciprocal teaching dialogue. (Contains 50 references.) (Author/RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED356456

Technical Report No. 574

RECIPROCAL TEACHING: A REVIEW OF 19 EXPERIMENTAL STUDIES

Barak Rosenshine
Carla E. Meister
University of Illinois at Urbana-Champaign

May 1993

Center for the Study of Reading

TECHNICAL REPORTS

College of Education
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
174 Children's Research Center
51 Gerty Drive
Champaign, Illinois 61820

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Fran Echr
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

CS011268

BEST COPY AVAILABLE

CENTER FOR THE STUDY OF READING

Technical Report No. 574

RECIPROCAL TEACHING: A REVIEW OF 19 EXPERIMENTAL STUDIES

Barak Rosenshine
Carla E. Meister
University of Illinois at Urbana-Champaign

May 1993

College of Education
University of Illinois at Urbana-Champaign
174 Children's Research Center
51 Gerty Drive
Champaign, Illinois 61820

1992-93 Editorial Advisory Board

Diane Bottomley

Eurydice M. Bouchereau

Clark A. Chinn

Judith Davidson

Colleen P. Gilrane

Heriberto Godina

Richard Henne

Carole Janisch

Christopher Kolár

Brian A. Levine

Elizabeth MacDonell

Montserrat Mir

Punyashloke Mishra

Jane Montes

Billie Jo Rylance

Shobha Sinha

Melinda A. Wright

MANAGING EDITOR
Fran Lehr

MANUSCRIPT PRODUCTION ASSISTANT
Delores Plowman

Abstract

Since the introduction of the concept of reciprocal teaching by Palincsar and Brown in 1984, researchers have conducted a number of studies that attempted to replicate and/or extend the original one. A review was conducted of 19 studies that used reciprocal teaching to help students improve their ability to comprehend text. The review found that (a) results were usually more significant when explicit instruction in the cognitive strategies was provided before the reciprocal teaching began than when reciprocal teaching only was used; (b) results were mostly non-significant when below-average students were taught, yet usually significant when all other students were taught; and (c) results were usually significant when experimenter-developed tests were used, yet usually non-significant when standardized tests were used.

RECIPROCAL TEACHING: A REVIEW OF 19 EXPERIMENTAL STUDIES

Following the introduction of the concept of reciprocal teaching in Palincsar and Brown's (1984) exemplary study, there have been a number of studies that attempted to replicate and/or extend the original one. Because of the theoretical and instructional interest, we prepared this report in which we review, organize, and summarize the results of all these studies and discuss the implications for both research and practice.

For this review, we searched the ERIC and Dissertation Abstracts International data bases for experimental studies that used reciprocal teaching. We selected studies if the words *reciprocal teaching* were explicitly stated and if the authors also referenced the work of Palincsar and Brown (1985) or the earlier work of Manzo (1969). Studies were selected only if they included both experimental and control groups and if they randomly assigned students to the two groups or determined the two groups to be similar on initial measures.

Defining Reciprocal Teaching

As is unfortunately common in education, terms are given a variety of meanings as they are used by researchers and teachers. Reciprocal teaching is no exception. As we gathered and read studies, we learned that two overlapping but different approaches to reciprocal teaching instruction had been used: the original Palincsar and Brown (1984) approach, which we labeled *reciprocal-teaching-only* (RTO), and a later but overlapping approach, which we labeled *explicit teaching before reciprocal teaching* (ET/RT).

Reciprocal-Teaching-Only (RTO)

In their classic study, Palincsar and Brown (1984) used the term *reciprocal teaching* to refer to a set of learning conditions in which children "first experience a particular set of cognitive activities in the presence of experts, and only gradually come to perform these functions by themselves" (p. 123). Students and teacher actively participate in a dialogue that focuses on a section of text being read, taking turns in leading the dialogue. During this dialogue, students learn to use four comprehension-fostering and comprehension-monitoring strategies: generating questions, summarizing, clarifying, and predicting.

The expert, or teacher, initially takes on the major responsibility for the group's activities. As described by Palincsar and Brown (1985),

> The teacher models and explains, relinquishing part of the task to novices only at the level each one is capable of negotiating at any one time. Increasingly, as the novice becomes more competent, the teacher increases her demands, requiring participation at a slightly more challenging level. (p. 13)

In addition to allowing active participation of students during the early stages of learning, RTO provides students with many opportunities to witness the success of the strategies in actual reading contexts (Palincsar & Brown, 1984), and provides the "teacher the opportunity to serve as on-line diagnostician for each child, providing him/her with appropriate instruction geared to the child's individual needs" (Lonberger, 1988, p. 17).

In the RTO approach, there is no explicit instruction on the individual strategies before the dialogues. During the dialogues, the teacher provides prompts, models, cues, and feedback on the use of cognitive strategies in the early stages of learning, as illustrated by the following examples taken from the Palincsar and Brown (1984) study:

(Questioning)

T: What would be a good question about pit vipers that starts with the word "why?"

S: (No response)

T: How about, "Why are the snakes called pit vipers?"

- - - - - - - - - -

T: That's good. Keep going.

S: How do spinner's mate is smaller than. . . How am I going to say that?

T: Take your time with it. You want to ask a question about spinner's mate and what he does, beginning with the word "how."

S: How do they spend most of his time sitting?

T: You're very close. The question would be "How does spinner's mate spend most of his time?" Now you ask it.

S: How does spinner's mate spend most of his time?

- - - - - - - - - -

S: Snakes' backbones can have as many as 300 vertebrae almost ___ times as many as humans."

T: Not a bad beginning, but I would consider that a question about a detail. See if the next time you can find a main idea question and begin your question with a question word--how, why, when. . . .

- - - - - - - - - -

(Summarizing)

T: That was a fine job, Ken, but I think there might be something to add to our summary. There is more information that I think we need to include. This paragraph is mostly about what?

S: The third method of artificial evaporation.

As the sessions continue, students in the group progressively take on more responsibility for carrying out the dialogue. The four strategies provide the students with a framework for discussing the text. Students in the group begin to provide models, hints, and prompts to each other, as well as feedback regarding the use of the strategies. Although the teacher is not absent from these later discussions, the role of the teacher shifts to that of a sympathetic coach.

We believe that 9 of the 19 studies in our review fit into the RTO category: Fischer Galbert, 1989; Helfeldt and Lalik, 1976; Jones, 1987; Labercane and Battle, 1987; Manzo, 1969; Padron, 1985; Palincsar and Brown, 1984; Rich, 1989; and Rush and Milburn, 1988 (see Appendix A).

Explicit Teaching Before Reciprocal Teaching (ET/RT)

As researchers conducted reciprocal teaching studies, another, overlapping form of reciprocal teaching emerged. This form of reciprocal teaching refers to the practice of providing explicit teaching of the cognitive strategies *before* the reciprocal teaching dialogues begin. An example of explicit teaching before reciprocal teaching (ET/RT) is the work of Palincsar (1987), who provided five days of teacher-led explicit instruction in the strategies of questioning, summarizing, predicting, and clarifying. This instruction--which involved teacher presentation, guided student practice, and independent student practice--took place *before* the reciprocal teaching dialogues began. (To illustrate what explicit instruction before the dialogues might look like, we include in Appendix C the instructional material that Palincsar [1987] developed and gave to the classroom teachers for teaching the strategy of question-generating.) In these studies, instruction took place both during the explicit teaching sessions and during the reciprocal teaching dialogues that *followed* the explicit teaching sessions. The purpose of this explicit instruction before reciprocal teaching was to introduce students to the strategies and accompanying vocabulary, not to ensure mastery of the four strategies. Teachers or experimenters provided prompts, suggestions, hints, explanations, feedback, and correction during the dialogues, as students applied the strategies to the discussion of the text. The amount of support provided was determined by the individual needs of students as they progressively took on more of the responsibility in the application of the strategies to the reading of text. As will be elaborated in the discussion, the ET/RT studies differed among themselves in both the specific cognitive strategies taught (e.g., summarization, question-generation, clarification, prediction) and in the prompts or cognitive facilitators (i.e., question type, signal words) used to teach these strategies.

Of the 19 studies we reviewed, 10 appeared to belong in the ET/RT category: Brady, 1990; Dermody, 1988; Levin, 1989; Lonberger, 1988; Lysynchuk, Pressley, and Vye, 1990; Nolte and Singer, 1985; Palincsar, 1987; Shortland-Jones, 1986; Taylor and Frye, 1988; and Williamson, 1989 (see Appendix B).

One reason for this wide range in instructional activities all taking place under the term *reciprocal teaching* is that the readers of the Palincsar and Brown (1984) article were not told precisely *how* cognitive strategies were taught during reciprocal teaching. Readers were given examples of dialogues illustrating how a teacher provided support during reciprocal teaching through the use of prompts and feedback. Unfortunately, the article contained little discussion of *how* the cognitive strategies were taught, or of whether specific prompts or facilitators were provided to help the students learn them. Because of this ambiguity, researchers conducting further studies varied in their interpretation of the concept of reciprocal teaching and varied in the type and amount of instruction that was built into their studies. (A related discussion on assessing teacher implementation of the instructional program can be found under the heading "Additional Assessments.")

One serendipitous advantage of such ambiguity is that there was a wide range of instruction in the subsequent studies, and we were able to probe these differences to learn new lessons about effective instruction. Given this natural variation, we hoped to learn more about reciprocal teaching by grouping studies according to the type of instructional approach used in the study, the type of student, and the type of outcome measure. We read and studied the instructional procedures used in each. Most of the ET/RT studies were easy to classify because they contained a description of the explicit teaching.

However, to be sure about our classification of studies, we wrote or phoned the authors--sometimes a second and third time. In instances where we could not locate the author of a dissertation, we contacted his or her adviser. All authors and advisers were fully cooperative, and we thank them. (It is unfortunate that there is no tradition of providing detailed descriptions of instructional procedures in publications. Indeed, these procedures are often printed in smaller type in many journals.

Summary

Our reading of these studies indicated that there were two approaches to reciprocal teaching: reciprocal-teaching-only, in which the instruction consisted of models, prompts, and hints and in which all the instruction took place during the dialogues; and explicit teaching before reciprocal teaching, in which instruction in cognitive strategies was provided *before* the dialogues began. To test whether these different approaches yielded different results, we analyzed these two types of studies separately.

Organizing Results

In addition to classification by *instructional approach*, we found that the studies could be classified by two other means: *type of student* and *type of outcome measure*.

Type of Student

Good-poor/all students. In about half the studies, the experimenters provided instruction to all the students in a class or grade level and included all of them in the analyses. We coded such studies as *all*. A second method of selecting students was to use information on their decoding and comprehension abilities. The original Palincsar and Brown (1984) study introduced the valuable idea of selecting students who were near grade level in decoding but below grade level in comprehension. We coded these studies as *good-poor*. Palincsar continued this practice in her 1987 study. This seems a particularly good approach because it enables one to select students having problems specifically with comprehension. Surprisingly, only 3 of the 15 studies that followed the Palincsar and Brown (1984) study used this practice (Dermody, 1988; Lysynchuk et al., 1990; Palincsar, 1987). Significant results on at least one outcome measure were obtained in all four studies that used good-poor students.

As our review progressed, we realized that only a few studies had selected good-poor students, and that the results of these studies were similar to the results obtained in the studies in which *all* the students in a classroom were instructed. Consequently, we combined the two groups into a classification we called *good-poor/all*.

Below-average students. The second group of studies involved students labeled *below-average*. This is a mixed group of studies, containing students in Chapter 1 classes (Fischer Galbert, 1989), learning-disabled classes (Labercane & Battle, 1987; Levin, 1989), and remedial summer classes (Manzo, 1969). This group also contained studies in which the authors stated that a majority of the students were reading a year or more below grade level (e.g., Brady, 1990).

Two of the studies included in this review involved teaching adults, and with some uncertainty, we also included them. Rush and Milburn (1988) used students in a postsecondary occupational training program in diesel mechanics, and we placed this study in the good-poor/all category. Rich (1989) taught adults studying for their General Education Degree (GED) certificate. The mean reading ability for this group was the fifth-grade level, but those students who were selected had an IQ score of 90 or above. We placed this study in the below-average classification.

Type of Outcome Measure

We created a third classification based upon the outcome measures used in the studies. Researchers used standardized reading tests or experimenter-developed tests, and some used both types of test. The experimenter-developed tests included passages followed by questions requiring short answers and/or requests to summarize material.

Thus, each study was classified three ways: (a) the type of instruction provided (RTO vs. ET/RT), (b) the type of students taught (good-poor/all vs. below-average), and (c) the outcome measures used (standardized reading tests vs. experimenter-developed tests). In this report, we present results separately for each classification.

Results

Overall Results

Table 1 presents results for all 19 studies, grouped according to whether the results were significant (S), mixed (S/NS) or non-significant (NS), and gives the overall median effect size. Each study was counted only once. In 5 studies that used two types of outcome measures (standardized tests and experimenter-developed tests), there were mixed results, that is, results were significant on either the standardized test or the experimenter-developed test, but not on both. We reported a single-effect size for these studies, and when listing significance of the results, we reported these studies as S/NS. Effect sizes were not computed for 6 studies in this review. In 5 of these studies, the researchers did not provide standard deviations, and in 1 study, the researcher did not provide the adjusted means.

Overall, the results are evenly split between significant and non-significant results with a median effect size of .57.

[Insert Table 1 about here.]

Additional Analyses

Table 2 presents results for all studies according to type of instructional approach, type of student, and type of test. Within each grouping, the results are classified as in Table 1.

[Insert Table 2 about here.]

Type of instructional approach. When we classified the results by type of instructional approach used--whether there was explicit instruction before beginning the reciprocal teaching dialogues (ET/RT) or whether there was reciprocal teaching only (RTO) without the prior explicit teaching--we found that the ET/RT approach was used in 10 studies, and RTO was used in 9 studies. Table 2 illustrates a trend that favors ET/RT (median effect sizes are .60 vs. .34). Five of the studies that taught cognitive strategies before the reciprocal teaching dialogues used two different outcome measures: a standardized test and an experimenter-developed comprehension test. Results were significant on one of the two tests in all of those studies. Overall, 9 of the 10 studies that explicitly taught cognitive strategies obtained significant results on at least one outcome measure. In contrast, when RTO was used, only 3 of the 9 studies obtained significant results. (Size of instructional groups and the duration of instruction are discussed under the heading of "Additional Results.")

Type of student. When we classified results by type of student, we found 12 studies that used good-poor/all students, and the majority of the results of these studies were significant. Only 7 studies were located in which below-average students were taught, and the majority of the results were non-significant.

The median effect size when good-poor/all students were taught was .57. The overall effect size was .48 for the below-average students. These results for the below-average students are somewhat skewed because we could only compute effect sizes for 4 studies, and also because we included the study by Rich (1989), which had an effect size of 1.44. If we were to exclude this study, the median effect size in this category of below-average students would drop dramatically to .14.

Type of outcome measure. Because two types of outcome measures were used in these studies, standardized tests and experimenter-developed tests, we analyzed the results in an attempt to discern any patterns produced by the type of test. Standardized tests were used in 13 studies, the Gates-MacGinitie Reading Test being the most frequently cited. A wide range of tests were found in the 11 studies that used experimenter-developed tests. These tests included experimenter-developed multiple-choice tests, short-answer tests, student summaries scored for total propositions recalled, and student summaries scored for the level of the ideas recalled. The two types of summaries were used most often. We included all of these outcomes under the category of experimenter-developed tests.

Five studies used both a standardized test *and* an experimenter-developed test: Brady, 1990; Dermody, 1988; Lysynchuk et al., 1990; Taylor and Frye, 1988; and Shortland-Jones, 1986. In the first 4 studies, the results were non-significant on the standardized test but significant on the experimenter-developed test. In the fifth study (Shortland-Jones, 1986), the results were reversed. Thus, when researchers used experimenter-developed tests, particularly requests to summarize, they obtained significant findings more frequently. (In Tables 2, 3, and 4, when presenting results by type of test, we included each of those 5 studies twice: once under results for standardized tests, and once under results for experimenter-developed comprehension measures.)

For studies that used standardized tests, 4 of the 13 studies obtained significant results (*es* = .32). The picture was different for experimenter-developed tests. Here, 8 of the 11 studies had significant results and 1 study had both significant and non-significant results (*es* = .87). Thus, these fairly strong results favor the use of experimenter developed tests.

Students x Test x Instructional Approach

For studies that provided instruction to the good-poor/all students (see Table 3), the majority of results were significant when experimenter-developed tests were used (*es* = .85), but not significant when standardized tests were used (*es* = .44). The largest effect sizes were obtained when both (a) cognitive strategies were explicitly taught before beginning the reciprocal teaching dialogues, and (b) experimenter-developed tests were used (*es* = .93).

The 2 studies that provided instruction to below-average students (see Table 4) and provided explicit teaching in cognitive strategies before the dialogues began (Brady, 1990; Levin, 1989) were successful on at least one outcome measure (*es* = .87). In contrast, of the 5 studies that did not provide explicit teaching before the dialogues, only 1 was successful (Rich, 1989) (*es* = .14). One should note that there are only 7 studies in this category. Thus, there is the beginning of a trend, albeit a shaky one, also favoring explicit teaching before reciprocal teaching for the below-average students.

[Insert Tables 3 and 4 about here.]

Type of test interactions. When we looked at the good-poor/all students by themselves (see Table 3) and the below-average students by themselves (see Table 4), no Type of Test x Type of Student interactions were found. For each type of student, the effect sizes were higher when researchers used experimenter-developed tests than when they used standardized tests (see Table 5).

[Insert Table 5 about here.]

Design differences by type of student. When teaching good-poor/all students, the most common practice was to explicitly teach cognitive strategies and to assess the results using both experimenter-developed tests and standardized tests. The results were significant in 6 of the 7 studies that did both (*es* = .93). In contrast, the most common practice when dealing with below-average students was (a) to use RTO and (b) to test the results using standardized tests. This practice was unsuccessful in 4 of 5 studies. Thus, although no one knew it at the time, from a design point of view, the below-average students appear to have been shortchanged; they received less explicit instruction in cognitive strategies, and researchers more frequently used standardized tests to assess their learning.

Additional Results

Grade Level

Instruction in reciprocal teaching was provided for a wide range of grade-level students, from 7-year-olds to adults. When we looked at results by grade level (see Table 6), we found both significant and non-significant effects at all grade levels, with the exception of third grade. Three third-grade classes in 3 separate studies (2 in the good-poor/all sample, and 1 in the below-average sample) yielded non-significant results. This issue needs to be studied further.

[Insert Table 6 about here.]

Number of Instructional Sessions

The number of instructional sessions provided in these studies ranged from 6 sessions to 100 sessions (See Appendixes A and B). We found no differences in the median number of sessions for significant, mixed, and non-significant studies across different types of students (see Table 7). For the good-poor/all students, both the significant and non-significant studies lasted for a median of 20 sessions, and the median was 17 for the studies with mixed results. For the below-average students, the median was 28 sessions for the two significant studies and 29 sessions for the four non-significant studies. Overall, there were more instructional sessions in studies involving below-average students, but again, more sessions were not associated with greater success.

We also looked at the number of sessions for studies in which ET/RT was used and those that used RTO. We found no differences in the total number of sessions for the two types of studies. Interestingly, the non-significant studies that used RTO had the highest number of median sessions.

[Insert Table 7 about here.]

Size of Instructional Group

Table 8 presents results by median size of the instructional group. Again, we found no relationship between the size of the instructional group and the significance of the results for either type of student.

Researchers obtained significant results with group sizes of 18 and 19 and non-significant results with group sizes of 4 and 5. If there is any pattern, it is for non-significant results to be associated with *smaller* group sizes, for both types of students.

[Insert Table 8 about here.]

Number and Types of Cognitive Strategies Taught

In 11 of the 19 studies, teachers provided instruction on the four comprehension-fostering strategies identified by Palincsar and Brown (1984). The remaining studies taught from one, two, or ten cognitive strategies (Levin, 1989). We found no relationship between the number of strategies taught and student achievement (see Table 9). Indeed, for the studies involving the good-poor/all students, the median number of strategies taught in the studies having significant results was two, and the median number of strategies taught in the studies having non-significant results was four. When below-average students were taught, no difference was found between the median number of strategies taught in those studies that significantly raised comprehension and those that did not. Interestingly, in 3 studies, only one strategy (question-generation) was taught (Helfeldt & Lalik, 1976; Manzo, 1969; Nolte & Singer, 1985), and 2 of these studies yielded significant results.

[Insert Table 9 about here.]

We found a pattern in the selection of specific cognitive strategies that were taught. If only one strategy was taught, that strategy was almost always question-generation. If two strategies were taught, the strategies were almost always question-generation and summarization. When four strategies were taught, they were question-generation, summarization, prediction, and clarification. There were only two exceptions to this pattern: one study taught graphic organizers as the third strategy (Shortland-Jones, 1986) and one study (Levin, 1989) taught the 10 Informed Strategies for Learning (ISL) developed by Paris, Cross, and Lipson (1984). Units of the ISL program incorporated the four strategies mentioned above and also taught other strategies, such as previewing the reading task to establish goals and purposes for reading, exploring different kinds of meaning, and making inferences.

Person Providing the Instruction

We were interested in whether different results occurred when studies were grouped according to the person providing the instruction--the experimenter or the teacher. These results are summarized in Table 10. We detected no differences. (A discussion of teacher/experimenter implementation can be found under the headings "Additional Assessments" and "Instructional Issues.")

[Insert Table 10 about here.]

Types of Control Groups

There were four types of control groups used in the studies, grouped according to the kind of instruction they received: basal, basal with direct instruction, read and answer questions, and other activities. A few studies had more than one control group. In those cases, we selected as *control* the group that had traditional basal instruction or that was more similar to traditional instruction. The following is an explanation of each type of control group.

Basal. This category refers to those studies in which the control group read stories and answered questions, but received no further instruction on how to construct answers to these questions. Eight of the studies were placed in this category.

Basal with direct instruction. We placed studies in this category when the control group received explicit instruction in reading comprehension, such as instruction in identifying the main idea of a paragraph (Lonberger, 1988), or in answering questions whose answers were explicitly or implicitly in the text (Padron, 1985; Palincsar & Brown, 1984). Six studies were placed in this category.

Thus, in 14 of the 19 studies, students in the control groups received traditional basal instruction or some form of direct instruction.

Read and answer questions. This category includes studies in which the control group silently read the same training passages used by the students in the experimental group. Two studies had control groups that fit this description. In 1 study (Rich, 1989), the students answered questions on the passages. Students in the second study (Lysynchuk et al., 1990) read the training passages, but we could not determine whether those students also answered questions.

Other. We placed in this category those studies that did not fit into any of the other categories. This category contains 3 such studies, 1 in which the control group did computer extension exercises not directly related to reading comprehension (Brady, 1990) but did take the daily comprehension tests, and 2 studies in which it was not stated what the control group did (Dermody, 1988; Labercane & Battle, 1987).

One might argue that only the first two types of control groups--basal and additional instruction--were the most appropriate control groups. Accordingly, we recomputed the overall results just for those 14 studies that used the first two types of control groups (see Table 11). When this was done, the overall pattern of significant and non-significant findings was unchanged. However, the overall effect size did drop from .57 to .40.

[Insert Table 11 about here.]

In summary, when we compared studies that had significant results with studies that had non-significant results, we found no differences in results by (a) grade level, (b) number of sessions, (c) size of the instructional group, (d) number of cognitive strategies that were taught, or (e) whether the experimenter or the teacher did the training. No differences were found in the pattern of results between S, S/NS, and NS studies when different control groups were considered, yet there was a slight decrease in the overall effect size when we compared only those studies in which the control groups received some form of instruction.

Additional Assessments

Implementation

In 11 of the studies, the experimenter conducted the lessons, and--with the exception of 1 study (Palincsar & Brown, 1984)--implementation of the reciprocal teaching procedure was not assessed. In the remaining 7 studies, experimenters trained classroom teachers to provide the instruction to the students, and the level of teacher implementation of the treatment was assessed in 6 of these studies (see Table 12). The level of implementation was reported as high in 4 of 5 successful studies. The level of implementation was also high in 1 unsuccessful study (Manzo, 1969). In a second unsuccessful study (Williamson, 1989), the researcher suggested that low teacher implementation of the reciprocal teaching approach may have contributed to the insignificant results obtained.

[Insert Table 12 about here.]

As shown in Table 12, these 7 studies differed in how the researchers assessed teacher implementation. For example, in the Manzo (1969) study, the frequency of questions asked by the teachers and students during the tutoring sessions and the type of questions asked (e.g., recall, translation, evaluation) served as the assessment; whereas, Levin (1989) observed teachers on 12 teaching behaviors (e.g., modeling and demonstration of the strategy, encouragement of metaphorical reasoning, use of praise and corrective feedback, and fading of support).

One possible reason for these differences is that Palincsar and Brown (1984) did not provide the implementation criteria. Although they describe how the teachers were trained and explain that "the first author checked weekly on the teacher-directed sessions to see if the intervention was being conducted properly" (p. 158), the authors give no information about how this decision was determined. Thus, subsequent researchers have not had specific implementation criteria to which they can refer, and so several researchers developed their own.

The original Palincsar and Brown studies did report the results of the students' daily comprehension assessments, and these assessments might serve as an indirect measure of implementation. However, similar daily comprehension assessments were used in only two studies, those by Lysynchuk et al. (1990) and Brady (1990).

In later articles, Palincsar and her associates (Palincsar, 1986; Palincsar, Stevens, & Gavelek, 1989) illustrated and made explicit the important instructional features of dialogues by comparing the dialogues of several primary teachers. However, none of the more recent studies in this review referenced these articles or referred to these critical features. (We include a list of these features in the discussion on quality of dialogue found under the heading "Suggestions for Classroom Practice and Future Research.")

Student Learning of the Strategies

In 8 of the studies reviewed, the researchers assessed the extent to which the students had learned the specific strategy. Four different assessments of student learning and the studies that used them are listed in Table 13. The results are discussed in this section.

[Insert Table 13 about here.]

Asking students to generate questions about a new passage. At the completion of the reciprocal teaching sessions in the Palincsar and Brown (1984) study, the researchers assessed student ability to generate questions. Students were given two passages (written at their seventh-grade level) and asked to "write 10 questions a classroom teacher might ask if testing a student's knowledge of a passage" (p. 134). The student questions were typewritten, corrected for grammar and spelling, and scored by two independent raters.

> They were asked to rate each question as a main idea question (worth two points) or a detail question (worth one point), as a question lifted directly from the text (zero points) or paraphrased (one point). In addition, the overall the quality of each question was rated on a five-point scale from one to five. Finally, a question that the rater indicated she would ask herself was awarded an extra point. (Palincsar & Brown, 1984, p. 150)

We present this assessment procedure in some detail because it is sensible, and it is one that might well be used in future research.

Surprisingly, although question-generation was one of the strategies taught in *all* of the studies, the ability to generate questions was assessed in only 7 studies: Dermody, 1988; Lysynchuk et al., 1990; Padron, 1985; Palincsar, 1987; Palincsar and Brown, 1984; Shortland-Jones, 1986; and Taylor and Frye, 1988.

The results of this assessment procedure are rather surprising. In 6 of the 7 studies, the students in the experimental group were *not* superior to students in the control group on the quality of the questions they wrote. Yet, despite this lack of a difference, students in 5 of those 6 studies were superior to control students on one of the two outcome measures. In other words, we found no relationship between ability to generate questions and reading comprehension scores. This relationship was also lacking in a study conducted by Risko and Feldman (1986), which found no relationship existed between the type of question asked about a story by three remedial second-grade students and their ability to answer that type of question on a posttest.

This is a perplexing finding. Here we have 6 studies in which the students in the experimental groups were superior to the controls on one of the major outcome measures, yet they were not superior on an implementation measure that assessed the learning of the strategy. In this case, we suggest, albeit tentatively, that in the successful studies that taught question-generation, the students in the experimental groups learned some additional processing procedure that they applied to the reading of new passages. The argument would be that what the students learned was not simply how to generate questions. Rather, as explained by Palincsar and Brown (1984), the new strategies they had acquired enabled the students to do deeper processing of what they read, to engage in making sense of what they read, to be aware of when they did not understand the material, and to engage in additional reading and searching when they encountered comprehension difficulties. These findings remain an interesting puzzle. We are not aware of this phenomenon occurring in other studies in which cognitive strategies such as study skills or summarization were taught.

Coding student questions during reciprocal teaching sessions. Another way of assessing student learning of the cognitive strategies is to assess their performance *during* the dialogues. As far as we could tell, such formal assessment took place in only 1 study (Manzo, 1969). In this study, 57% of the 145 questions generated by the students in the experimental group were questions requiring translation, conjecture, and explanation. In contrast, only six questions asked by students in the control group were similarly coded. Interestingly, despite these differences in questioning, there were no significant differences between experimental and control students' scores on the standardized reading tests.

Asking students to summarize a passage. In 6 of the studies, after completing the strategy training, students were asked to read a new passage and to write a summary of it (Dermody, 1988; Lysynchuk et al., 1990; Palincsar, 1987; Palincsar & Brown, 1984; Shortland-Jones, 1986; Taylor & Frye, 1988). The summary was then rated, usually by scoring each idea for its importance in the passage. Significant differences favoring the experimental group were obtained in 4 studies (Dermody, 1988; Palincsar, 1987; Palincsar & Brown, 1984; Taylor & Frye, 1988) but not in the fifth (Shortland-Jones, 1986). In the sixth study (Lysynchuk et al., 1990), one grade had significant results and one grade did not.

We suggest that a measure of student ability to summarize is both a measure of implementation and also a legitimate comprehension outcome measure; therefore, we also used summarization as an outcome measure in the previous section.

Asking students to make predictions. Although students were taught to make predictions in 12 of the 18 studies, they were assessed on this ability in only one study. Dermody (1988) asked students to make predictions on what the story might be about by simply reading the title of the story. The results

significantly favored the experimental group over the control group on ability to make predictions. (Concerns regarding this particular strategy, as expressed by Brady [1990], are discussed under the heading "Differences in Selecting and Teaching the Cognitive Strategies.")

Summary. Four procedures were used to assess student learning of the strategies, but only two appeared more than once: student ability to generate questions and student ability to summarize a portion of text. Interestingly, in most of the studies there was no difference between experimental and control students in ability to generate questions; nonetheless, the experimental students were usually superior on the comprehension measure. We interpret this result by suggesting that the students in the experimental group learned some additional processing, processing that was not assessed simply by asking students to generate questions.

Instructional Issues

A Comparison of Different Settings for Teaching Cognitive Strategies

In the studies we reviewed, instruction occurred in two settings: explicit teaching before beginning the dialogues and instruction during the dialogues. By and large, researchers obtained significant gains more often in settings that provided explicit teaching before beginning the dialogues.

A possible third instructional approach for improving student comprehension of text is that of providing instruction without reciprocal teaching dialogue or cooperative learning. Five of these studies are described below.

Bereiter and Bird (1985) identified strategies used by expert readers when they had difficulty comprehending text (for example, rephrasing the material in simpler terms or reforming the difficulty into a problem). They taught these strategies to readers by first defining the strategies, reading passages aloud, and demonstrating the strategies as they thought aloud. The next step consisted of the instructor reading new passages aloud, applying the strategies, and asking the students to identify the strategy being used. Finally, the students read passages and orally marked their use of each strategy. A similar instructional procedure was used by Raphael and Pearson (1985) as they taught students to generate different types of questions. The researchers first modeled the questions and identified their type. Next, they modeled the questions, and students identified their types. Finally the students generated the questions and identified their types. Davey and McBride (1986) taught students to generate questions using prompts, models, guided practice, and a self-evaluation checklist. Students worked alone generating questions on passages and applying the self-evaluation checklist. In a study by Cohen (1983), students were provided with models and guided practice and then worked alone to complete "question-generating booklets." Blaha (1979) provided students with three questions that helped them to summarize. The students were first taught in a group to generate and answer these questions, and then individually practiced generating and answering questions. Before each session ended, the students came together in a group to compare answers and clear up difficulties.

In all five studies, students in the experimental group had significantly higher posttest scores than students in the control group on both a standardized reading test and experimenter-developed comprehension tests.

Differences in Selecting and Teaching the Cognitive Strategies

Type and number of strategies taught. In the 19 reciprocal teaching studies we reviewed, experimenters taught from one to ten cognitive strategies. No relationship existed between the number of strategies

taught and whether the results of the study were significant, not significant, or mixed (see Table 9). Studies that taught only one strategy, four strategies, or ten cognitive strategies all produced significant gains in comprehension.

However, we did detect a pattern in the selection and combination of the cognitive strategies taught. If only one strategy was taught, that strategy was almost always question-generation. If two strategies were taught, the strategies were usually question-generation and summarization. If four strategies were taught, they included the same four introduced in Palincsar and Brown's reciprocal teaching study: prediction, question-generation, summarization, and clarification. We were surprised that teaching one strategy appeared to be as effective in improving reading comprehension as teaching two or four strategies. This result raises questions about the internal processing that each strategy generates and about the new reading approach that might develop as a result of applying the strategy.

Brady (1990), for example, noted that when students used a clarification activity, it was most often in the form of requests for definitions. Even when using context, the glossary, and classmates, students were able to produce an appropriate definition only some of the time. Brady, therefore, was unconvinced of the efficacy of this activity. On the other hand, he did note that the effective result of this activity was to legitimize students' efforts to seek help from and offer assistance to each other, and that this might be part of what students liked about reciprocal teaching.

Brady also noted that the prediction activity was seldom successful. In his study, he found that the social studies text the students were using made it "virtually impossible to accurately predict the content of the next paragraph" (pp. 104-105). This was partly because the text lacked coherence and partly because of the chronological nature of the text. Brady recommended that future studies should evaluate the effectiveness of this activity through the use of both poorly and well-organized text.

Resnick (1987) agreed that we know little about how training procedures operate to produce their effects. That is, how can instruction in strategies such as question-generation or summarization, which may not be the actual components used by skilled readers, improve automatic comprehension processes? Researchers are not sure how the processing works. Perhaps these strategies are used by skilled readers but in an invisible form. Perhaps these strategies are only used when smooth comprehension breaks down. Perhaps practice in deliberate searching for meaning helps students develop approaches ("knowledge structures") that are applied to subsequent reading. Resnick stated:

> Research has located a "psychological space" in which educationally powerful effects seem to occur, but it has not yet adequately explained what happens in that space to produce the effects. Until we can produce a more substantial explanation, . . . it will be difficult to determine in advance the essential components of a training approach. (p. 27)

Differences in the concrete prompts used. A key instructional procedure in the studies that used explicit teaching was to provide students with specific, concrete prompts or "procedural facilitators" (Scardamalia & Bereiter, 1985) that served as aids students could refer to as they learned the cognitive strategy. We identified both the different prompts used in these studies and the procedures used to teach them (see Table 14).

[Insert Table 14 about here.]

Prompts for question-generation. Four different concrete prompts were used in 10 studies to help students learn the strategy of question-generation. The most common concrete prompt, used in 6 studies, provided students with signal words such as *when*, *why*, and *how* (Brady, 1990; Lysynchuk et al., 1990; Palincsar, 1987; Palincsar & Brown, 1984; Taylor & Frye, 1988; Williamson, 1989). A second

prompt, question types, was used to show students that there were three types of questions--those with answers that lay in a single sentence, those with answers that required combining information, and those with answers that lay outside the text. Students then used this concept to help them generate questions (Dermody, 1988). The third prompt, locating main ideas, instructed students to use the main idea as an aid in generating questions (Lonberger, 1988; Rich, 1989), and the fourth prompt consisted of teaching students a story grammar (story elements such as setting, character, and problem) to aid in generating questions (Nolte & Singer, 1985). Each concrete prompt was used in at least 1 study that obtained significant results. Unfortunately, there were too few studies using each type of prompt to enable us to discuss whether any one prompt was more successful than the others.

Prompts for summarization. Two major concrete prompts were used in 5 studies to assist the learning of the cognitive strategy of summarization. Both Palincsar (1987) and Brady (1990) taught students how to apply a set of procedures that were derived from the procedures used by experts to summarize, procedures first identified by Kintsch & van Dijk (1976).

Step 1. Identify the topic sentence

Step 2. If the topic sentence is not given, then identify the topic and the most important information about that topic.

Rule 1. Leave out unimportant information.

Rule 2. Give steps or list a title.

Rule 3. Cross out information that is redundant/repeated.

Brady added an additional prompt, providing the students with a sentence starter: "This paragraph tells us that. . . ." Brady explained that this additional aid was developed during the pilot testing to avoid a frequent error made by students who often made too general a summary statement such as "This paragraph was about horses" or "This paragraph is about herds."

Other researchers (Dermody, 1988; Lonberger, 1988; Taylor & Frye, 1988) taught students a "legs-and-table" procedure. Students were taught to list the main details in a paragraph and then use these details as a prompt to generate the summary.

The two concrete prompts used to help teach summarization, the Kintsch and van Dijk procedure and the legs-and-table procedure, have a number of separate parts or components. Each component was presented and practiced separately. First, the teacher presented one step or component, such as "leave out unimportant information," and then provided students with practice in applying that component to new examples. After each step or component was learned and practiced separately, students consolidated the components and applied them as a complete strategy. In summary, a form of subskills--we might call them cognitive subskills--appeared in the teaching of cognitive strategies for summarization.

Both of the summarization prompts discussed in this section were found in successful studies. Four other studies included summarization in the reciprocal teaching dialogues but did not describe or make explicit any prompts for teaching summarization (Levin, 1989; Lysynchuk et al., 1990; Palincsar & Brown, 1984; Williamson, 1989). Two of these studies obtained significant results (Levin, 1989; Palincsar & Brown, 1984), 1 obtained mixed results (Lysynchuk et al., 1990), and 1 obtained non-significant results (Williamson, 1989).

In sum, concrete prompts or procedural facilitators are important for helping students learn cognitive strategies. In these studies, a variety of concrete prompts successfully supported students as they learned the cognitive strategies. It is possible that different concrete prompts can be equally successful, but there is a need for further study on the characteristics of successful prompts.

Suggestions for Classroom Practice and Future Research

Quality of the dialogue. Dialogue plays a critical role in providing scaffolded instruction (Palincsar, 1986). Palincsar and her associates (Palincsar, 1986; Palincsar et al., 1989) attempted to describe the important instructional features of dialogues by comparing the dialogues of several primary teachers. They found that the less successful teachers often provided mere recitation of information about the strategy being taught during the dialogues. Students' involvement occurred mostly at the word level as they were asked to provide labels for strategies or to complete sentences begun by the teacher. In contrast, the more successful teachers engaged students more at the *idea* level and focused the dialogue on using the strategies as part of the discussion of the passage. They encouraged students to engage in the dialogue by rephrasing a question when students were unable to respond and by using and elaborating upon a student's response. They provided support by explicitly modeling ("If I were asking a question, I . . ."; "If I were going to summarize, I . . .") and helped students label the language or features of the interaction.

The following features, which emerged in Palincsar's (1986) analysis as distinguishing certain dialogues, might be used to evaluate implementation of reciprocal teaching in future studies:

1. teacher support of the students' contribution to the dialogue was at the idea level as opposed to the word level;

2. there was deft use of student ideas and linking of those ideas to new knowledge;

3. the dialogues had focus and direction;

4. the point of instruction was made explicit to the student and seemed explicit to the teacher; and

5. evaluative statements were made in ways that changed the complexion of a student response from negative to constructive. (p. 96)

Teaching the strategies. As noted, the more successful studies also provided direct instruction in the cognitive strategies before the dialogue began. To help in future studies, and in classroom practice, we present a few interesting instructional procedures for teaching the strategies that appeared in some of these studies.

1. When concrete prompts have a number of meaningful components, teach each component separately, consolidating them in the end. The Kintsch and van Dijk procedure and the legs-and-table procedure, which were used to teach summarization, have a number of separate parts or components, and each component was taught separately (Palincsar, 1987).

2. Regulate the difficulty of the materials for the students by selecting practice materials that are below the grade level of the students (Levin, 1989; Lonberger, 1988; Rich, 1989).

3. Regulate the difficulty of the instructional task by starting with simpler tasks such as generating questions about a single sentence and then moving to questions about a paragraph (Brady, 1990; Palincsar, 1987).

4. Provide cue cards that contain the concrete prompts to which students can refer during practice (Brady, 1990).

5. Provide self-checking or fix-up procedures to which students can refer after completing the task (Lonberger, 1988).

6. Gradually increase student responsibility, supporting students as they perform certain aspects of the task (Lonberger, 1988; Nolte & Singer, 1985).

It was unfortunate that out of the 19 reciprocal teaching studies, only 1 (Brady, 1990) discussed the problems that emerged during the dialogues. Such information would be useful for both teachers and researchers implementing reciprocal teaching and would contribute greatly to our collective knowledge on classroom instruction. We hope it will appear in future studies.

Assessing implementation. One of the shortcomings of this research and of the dissemination of reciprocal teaching is that little attention has been paid to implementation. Those who wish to implement reciprocal teaching in their classrooms or who wish to conduct future studies do not have a source they can turn to for guidance in practice and for assessment of implementation.

We hope that the above sections and our listing of the implementation procedures in Table 12 contribute toward that development, but there is a need for increased work not only in developing assessment procedures but also in identifying and discussing problems that occur during the instruction and the dialogues.

Shticklach

This section contains three topics that did not readily fit in the above discussions of assessments or instruction: (a) significant results with below-average students, (b) unexpected findings, and (c) the exemplary contributions of Palincsar and Brown (1984) and Brady (1990).

Significant Results with Below-average Students

We remain impressed with the selection criteria used by Palincsar and Brown (1984) in their original pioneering study--the selection of those students who were near grade level in decoding but below grade level in comprehension. This selective approach enables one to place poor readers in their "area of sensitivity to instruction." These selection criteria imply that reciprocal teaching is not likely to be successful with students who are poor in decoding. Yet, significant results were obtained in 3 of the 7 studies that provided instruction to below-average students or poorer readers. However, sufficient data were not provided in these studies to determine the decoding ability of the below-average students. The 3 studies are described below.

The subjects in Brady's (1990) study were 6 Native Alaskan students who had reading comprehension test scores below the 50th percentile--3 with scores two or more years below grade level. Brady conducted separate analyses by type of test item and found that the experimental students did quite well.

They were superior to control students not only on text-explicit items in the comprehension tests but also on the inferential text-implicit items and on the items for which background knowledge was required. The test materials, those developed by Palincsar (1987), were at the grade level of the students.

In Brady's study, the students used their own social studies and science textbooks, and completed pretests and posttests based on seventh-grade material (the students' grade level). If Brady's materials were at or near grade level, then his students did exceptionally well in an area beyond their expected competence as determined by standardized reading tests. As noted, his students also did well on test items for which outside background knowledge was required. Brady stated that although his students were poor decoders, they could still "plug their way through the texts that they had to read." He explained further:

> The continual emphasis on making meaning that is required by the questioning, summarizing, and predicting activities actually forced the students to move beyond a belief that decoding the words was a sufficient response to a request that they "read" a passage. This emphasis on making meaning pushed them through the decoding and into making sense of what they read. I felt that for some of them, this was the first time the concept of reading included this element as a major concern. (personal communication, September 26, 1991)

In the study by Levin (1989), which also obtained significant results, the students were sixth-, seventh-, and eighth-grade special education minority students, all of whom had been classified as learning disabled. There were a number of interesting elements in the Levin study. First, the instructional period was 10 weeks long. Each week the students learned another reading strategy. On the first two days of the week, the teacher taught the strategy directly. On the third day, the teacher provided students with guided practice. On the last two days, the students provided examples for each other. Levin made two decisions that may have contributed to the significant outcomes. First, she prepared instructional materials at both the third- and fifth-grade levels "so that the skill could be introduced with less challenging reading, and later applied to more difficult material" (p. 75). Second, the outcome measure was the comprehension scale of the third-grade level of the Stanford Achievement Test, scored for out-of-level students.

Five studies provided instruction for below-average students throughout the reciprocal teaching dialogue only. Only one of those studies was successful. In that study (Rich, 1989), the students were adult poor readers taking a course that prepared them for the high school equivalency exam, that is, the GED. Reading levels for these students ranged from 3.7 to 7.1. The mean IQ of these students was 93.2, with a standard deviation of 3.6. Rich used reading selections taken from fifth- to seventh-grade science textbooks as the instructional material, and the reading selections for the pretest and posttest came from reading material with a grade level of 6.7 (Rich, 1989, pp. 54-56).

In both the Levin and the Rich studies, the instructional materials and the test materials were *below* the students' grade levels and closer to their reading levels, thus situating the instruction within the students' zone of sensitivity to instruction (Palincsar & Brown, 1984). In general, this procedure of using instructional material closer to the actual reading level of the students should be considered for future practice.

In summary, these studies offer two suggestions that might be considered when preparing instruction for students who are poor in both decoding and comprehension: (a) provide for explicit teaching in the cognitive strategies, and (b) use reading materials that are close to the students' reading level.

Unexpected Findings

Some unexpected findings previously uncovered and discussed in this review of reciprocal teaching studies deserve repeated emphasis.

1. Although Palincsar and Brown (1984) inferred that it is important to situate the teaching of the cognitive strategies in the context of the dialogues, the results in our review suggest that significant student learning occurred when cognitive strategies were taught *before* the dialogues began.

2. Although it has been suggested that cognitive strategies are best taught holistically (Pearson & Dole, 1987; Tharp & Gallimore, 1988), in these studies the strategy of summarizing was frequently and successfully taught through the use of procedural facilitators that divided the strategy into subgoals. After successful completion of each subgoal, they were consolidated and practiced as a complete strategy.

3. We were surprised to find no difference in the results associated with the number of instructional sessions (see Table 7), even among the explicit teaching (ET/RT) studies. Nor were there differences in results associated with the number of students in the instructional groups (see Table 8), or the specific strategies taught, or the number of strategies taught (see Table 9).

The Contributions of Palincsar and Brown (1984) and Brady (1990)

Palincsar and Brown (1984) have made a number of major contributions to the teaching of reading. The first is the introduction of the terms comprehension fostering and comprehension monitoring. Such terms place a greater emphasis upon teaching the students skills that they can use to enhance their comprehension. These terms help provide a shift from the teacher monitoring comprehension by asking students questions about a story to providing students with activities they can use independently to monitor and enhance their own comprehension.

A second contribution is the identification of four such comprehension-fostering and comprehension-monitoring activities: self-questioning, summarizing, clarifying, and predicting. The combination of only four pervasive strategies is a great improvement over previous listings of 150 "reading skills" that appeared in many basal readers.

A third contribution is Palincsar and Brown's refinement and popularization of the instructional concepts of reciprocal teaching and scaffolding. That contribution, and the studies that followed their work, allow us a closer examination of various instructional approaches.

The fourth contribution is the assessment of student learning *during* the study. These assessments, which were discussed more fully in the section on comprehension, included (a) assessment of student ability to generate questions, (b) assessment of student ability to summarize, and (c) monitoring of student comprehension (comprehension probes) throughout the training. These assessments, which seldom occurred in instructional intervention studies, serve to advance the state of the art when doing future intervention studies. However, we were surprised to see that these design elements, clearly stated in the Palincsar and Brown (1984) study, appeared in only 5 of the 13 studies that followed Palincsar and Brown. The assessment of student comprehension throughout the intervention and the assessment of whether students learned the strategies would seem important in studies that taught cognitive strategies.

The final contribution made by Palincsar and Brown is the criteria for selection of students: Only students near grade level in decoding but below grade level in comprehension were selected.

Four contributions by Brady (1990) to the design of studies also merit mention. First, Brady provided students with cue cards containing specific prompts on question-generation and summarization to use during the dialogues. Second, he presented data for all three types of questions on the comprehension tests. Third, he conducted follow-up testing 60 days after the posttest, and fourth, he presented a clear discussion of problems faced during the dialogues.

Summary of All Results

1. Reciprocal teaching overall. The overall results for all studies that used reciprocal teaching, regardless of instructional approach, are evenly distributed between significant and non-significant results with an overall median effect size of .57 (see Table 1).

2. Instructional approach. The results favor ET/RT (*es* = .60, Table 2) rather than RTO (*es* = .34, Table 2).

3. Type of student. Overall, the results are equally effective for the two types of students, with a median effect size of .57 (see Table 3) for the studies using good-poor/all students and .48 (see Table 4) for the studies using below-average students. However, as we shall see below, there are some interesting "interactions" between students and type of test.

4. Type of test. There was a strong effect favoring experimenter-developed tests. Across all students the effect size was .32 (see Table 5) when standardized reading tests were used and .87 (see Table 2) when experimenter-developed tests were used.

5. Instructional approach by type of student. In studies providing instruction to good-poor/all students (see Table 3), reciprocal teaching was more effective when experimenter-developed tests were used (*es* = .85) than when standardized reading tests were used (*es* = .44). With these students, the strongest effects were obtained when there was the combination of teaching cognitive strategies before reciprocal teaching *and* using experimenter-developed tests (*es* = .93). For the below-average students (see Table 4), there was only a single study in which cognitive strategies were taught with ET/RT and experimenter-developed tests were used, and it was successful (*es* = .87).

For studies providing instruction to below-average students (see Table 4), effect sizes were higher following ET/RT (*es* = .87) than following RTO (*es* = .14). Effect sizes were much higher when experimenter-developed tests were used (*es* = 1.16) than when standardized reading tests were used (*es* = .14). The *weakest* effects were obtained when the RTO instructional approach was used and results assessed using standardized reading tests (*es* = -.21).

In sum, reciprocal teaching was most effective, for both types of students, when ET/RT and experimenter-developed tests were used. The results were particularly weak when RTO was provided for below-average students and standardized reading tests were used.

6. Grade level of students. Students of all ages, 7 years through adult, were found in both significant and non-significant studies. However, results were not significant in the 2 studies (Jones, 1987; Williamson, 1989) that instructed only third-grade students (see Table 6).

7. Number of sessions. The number of sessions ranged from 6 to 100. No relationship was found between the number of sessions and the significance of the results. No difference was detected when we compared the number of sessions between studies that provided ET/RT and studies that provided RTO (see Table 7).

8. Size of instructional group. Instructional group size ranged from 1 to 22 students. We investigated the relationship between the size of the instructional group and success. Surprisingly, for both the good-poor/all students and the below-average students, the median group size for the significant studies was *larger* than the median group size for the non-significant studies (see Table 8).

9. Number of strategies taught. The number of strategies taught ranged from one to ten; however, in 11 of the studies, the four cognitive strategies of prediction, clarification, question-generation, and summarization were taught. No relationship was found between number of strategies taught and student achievement. There were successful studies in which only one cognitive strategy was taught (question-generation) and a successful study in which the ten cognitive strategies of ISL (Paris et al., 1984) were taught (see Table 9).

10. Experimenter versus teacher. We also classified studies according to whether the experimenter or the classroom teacher provided the instruction. There were 9 studies in each category, and the results were identical for both. In each category, 3 of the studies had significant results, 3 had both significant and non-significant results, and 3 had results that were non-significant (see Table 10).

11. Type of control group. When we recomputed the overall results for only those studies that used basal or additional instruction as the control group ($N=14$), the overall pattern of significant and non-significant findings was unchanged. However, the overall effect size dropped from .57 to .40 (see Table 11).

Conclusion

Six interesting questions emerged from our analysis of this research, and our answers to them may be of interest to both teachers and researchers.

How effective is reciprocal teaching in improving student comprehension?

Reciprocal teaching was more effective when teachers provided explicit teaching of cognitive strategies before the reciprocal teaching dialogue. This trend was strongest in studies using the combination of good-poor/all students and experimenter-developed tests. Only a few studies involved below-average students, but the trend for these few studies favors providing below-average students with explicit teaching before the dialogues and *not* using standardized reading tests.

Among all types of students, significant results occurred more often when experimenter-developed tests were used, and less frequently when standardized tests were used.

The results were weakest when below-average students (those reading well below grade level or those found in remedial or learning disabled classrooms) were involved in RTO and, particularly, when researchers used standardized tests to assess the outcomes.

What are the more effective cognitive strategies to teach?

Most of the studies taught the four strategies introduced by Palincsar and Brown (1984), strategies that serve to promote both comprehension-fostering and comprehension-monitoring activities. Other studies that obtained significant results taught only the strategy of question-generation, the combination of

question-generation and summarization, or the 10 ISL strategies (Levin, 1989). Based on the studies in this review, one cannot make a definitive statement on which strategies are more effective; there is a need for more research on the effects of individual strategies and combinations of strategies. *Nevertheless, based on this research, the strategy of question-generation, taught individually and in* combination with other strategies, had the highest success rate, followed by summarization. In the future we would be particularly interested in studies contrasting the learning of question-generation and summarization where the results are assessed by both summarization and another type of comprehension measure.

More research is needed on the cognitive strategies of prediction and clarification, particularly when contrasted with question-generation and/or summarization. Brady (1990) suggested that it is difficult for students to use prediction as a strategy when they read history textbooks because these textbooks use a chronological sequence rather than a canonical form. Brady also noted that the clarification strategy is difficult for students to use when they read history textbooks because the new terms they encounter in these textbooks are frequently specific to the text. We would be interested in further exploration of Brady's ideas.

Strategies similar to clarification are the "fix-up strategies" or procedures one activates when comprehension difficulties arise while reading a text. Bereiter and Bird (1985) identified four main on-line comprehension strategies used by expert readers--restating material, rereading, demanding relationships between text information, and formulating the difficulty as a problem--and demonstrated that teaching such strategies can lead to improved comprehension. We would be interested in seeing these fix-up strategies taught through reciprocal teaching.

In sum, we are uncertain which are the more effective strategies to teach with reciprocal teaching or how many strategies to teach. The two strongest candidates are question-generation and summarization, both of which serve a comprehension-fostering function in that they require a student to search the text and perform deeper processing. They also serve a comprehension-monitoring function in that difficulty in performing either task signals the student that there are comprehension difficulties.

What are the more conducive settings for teaching cognitive strategies?

Is it more effective to explicitly teach cognitive strategies before beginning the reciprocal teaching, or are such strategies best introduced and taught in the context of the reciprocal teaching dialogues? In this review, although there were successful studies using each approach, the balance of significant to non-significant results and the effect sizes favor providing students with *explicit teaching* in cognitive strategies *before* the reciprocal teaching dialogue begins. This finding is particularly strong when the good-poor/all students are taught and the results assessed using experimenter-developed tests. Only 3 successful studies provided instruction to below-average readers, and 2 of these studies provided ET/RT. Good models of studies that used ET/RT are those conducted by Brady (1990), Palincsar (1987), and Taylor and Frye (1988).

Based on this research, we would recommend the ET/RT approach because it has been successful with both regular and below-average students.

We also note a third possible approach; other studies have successfully taught question-generation or summarization using explicit instruction procedures but without using reciprocal teaching (Bereiter & Bird, 1985; Davey & McBride, 1986; Grajia, 1988). Future studies might contrast this approach, explicit instruction without reciprocal teaching, with RTO to gain further insights into the effects of reciprocal teaching.

What are the more effective prompts for teaching these strategies?

We looked at the facilitators or prompts that were used to teach the two major cognitive strategies--question-generation and summarization--in hopes of identifying which were the more effective. We were somewhat surprised to find that in the context of these studies, all of the prompts used to teach question-generation and summarization were equally successful. Indeed, 9 of the 10 studies that taught cognitive strategies using concrete prompts were successful in at least one of the two types of outcome measures. The question, then, of which facilitators are more useful for teaching cognitive strategies awaits future experimental studies that contrast the use of different facilitators for teaching the same cognitive strategy.

What instructional procedures are useful for teaching the strategies?

To answer this question, we read through the instructional procedures of each study in an attempt to identify interesting variables. We found six such variables:

1. regulate the difficulty of the materials by starting with materials below the grade level of the students.

2. regulate the difficulty of the instructional task by starting with the teacher modeling the first part of the task.

3. provide cue cards giving the prompts (i.e., teacher-generated questions or question-signal words), which students can refer to during the dialogues.

4. model the process of using the procedure in developing questions and summaries.

5. provide models of good questions and summaries.

6. increase student responsibility during the dialogues by gradually diminishing the prompts and models, increasing the complexity of the material, and putting all the component parts or steps together.

What outcome measures produced more significant results?

Two types of outcome measures were used in these studies: standardized reading tests and experimenter-developed tests. Four of the 13 studies that used standardized reading tests to assess comprehension had significant results, compared to 8 studies out of 10 having significant results on experimenter-developed tests. Further, in 4 of the 5 studies that used both forms of tests, the results were significant when experimenter-developed tests were used and non-significant when the standardized tests were used. It appears, then, that significant results were obtained much more frequently when experimenter-developed tests were used. We are not sure why this difference exists between standardized tests and experimenter-developed tests. One probable explanation is that the options offered as possible answers on the standardized tests were more discriminating, whereas most of the experimenter-developed tests required summarization or short answers to questions about a passage and did not distract or confuse students with highly discriminating options from which to choose the best answer.

For which type of students is teaching these strategies more effective?

The procedure of providing explicit teaching before the reciprocal teaching dialogues appears to have been effective for all types of students, from fourth graders through adults. For the good-poor/all

students, 7 of 8 studies yielded significant results on at least one dependent measure when explicit teaching was used. This pattern, favoring explicit teaching, held even when we removed those studies instructing students who were good in decoding but poor in comprehension and tallied only those studies using all students. In contrast, only 2 of 4 studies yielded significant results when RTO occurred. RTO was particularly unsuccessful when below-average readers were taught without prior assessment of whether the problem was in decoding or in comprehension. With the below-average students, the 2 studies that provided explicit teaching were successful. In contrast, only 1 of the 5 studies that used RTO was successful. Note, however, only a limited number of studies with below-average students was available for this review.

References

Bereiter, C., & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. *Cognition and Instruction, 2,* 131-156.

Blaha, B. A. (1979). *The effects of answering self-generated questions on reading.* Unpublished doctoral dissertation, Boston University.

Brady, P. L. (1990). *Improving the reading comprehension of middle school students through reciprocal teaching and semantic mapping strategies.* Unpublished doctoral dissertation, University of Oregon, Eugene.

Cohen, R. (1983). Self-generated questions as an aid to reading comprehension. *The Reading Teacher, 36,* 770-775.

Davey, B., & McBride, S. (1986). Effects of question-generation training on reading comprehension. *Journal of Educational Psychology, 78,* 256-262.

Dermody, M. (1988, February). *Metacognitive strategies for development of reading comprehension for younger children.* Paper presented at the annual meeting of the American Association of Colleges for Teacher Education, New Orleans, LA.

Fischer Galbert, J. L. (1989). *An experimental study of reciprocal teaching of expository text with third, fourth, and fifth grade students enrolled in Chapter 1 Reading.* Unpublished doctoral dissertation, Ball State University, Muncie, IN.

Grajia, M. L. (1988). *Direct instruction of a summarization strategy: Effect on text comprehension and recall in learning disabled students.* Unpublished doctoral dissertation, The Pennsylvania State University, University Park.

Helfeldt, J. P., & Lalik, R. (1976). Reciprocal student-teacher questioning. *Reading Teacher, 33,* 283-287.

Jones, M. P. (1987). *Effects of reciprocal teaching method on third graders' decoding and comprehension abilities.* Unpublished doctoral dissertation, Texas A&M University, College Station.

Kintsch, W., & van Dijk, T. A. (1976). Toward a model of text comprehension and production. *Psychological Review, 85,* 363-394.

Labercane, G., & Battle, J. (1987). Cognitive processing strategies, self-esteem, and reading comprehension of learning disabled students. *Journal of Special Education, 11,* 167-185.

Levin, M. C. (1989). *An experimental investigation of reciprocal teaching and informed strategies for learning taught to learning-disabled intermediate school learners.* Unpublished doctoral dissertation, Teachers College, Columbia University, New York.

Lonberger, R. B. (1988). *The effects of training in a self-generated learning strategy on the prose processing abilities of fourth and sixth graders.* Unpublished doctoral dissertation, State University of New York at Buffalo.

Lysynchuk, L. M., Pressley, M., & Vye, N. J. (1990). Reciprocal teaching improves standardized reading-comprehension performance in poor comprehenders. *Elementary School Journal, 90,* 471-484.

Manzo, A. V. (1969). *Improving reading comprehension through reciprocal questioning.* Unpublished doctoral dissertation, Syracuse University.

Nolte, R. Y., & Singer, H. (1985). Active comprehension: Teaching a process of reading comprehension and its effects on reading achievement. *The Reading Teacher, 39,* 24-31.

Padron, Y. N. (1985). *Utilizing cognitive reading strategies to improve English reading comprehension of Spanish-speaking bilingual students.* Unpublished doctoral dissertation, University of Houston.

Palincsar, A. S. (1982). *Improving the reading comprehension of junior high students through the reciprocal teaching of comprehension-monitoring strategies.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Palincsar, A. S. (1986). The role of dialogue in providing scaffolded instruction. *Educational Psychologist, 21,* 73-98.

Palincsar, A. S. (1987, April). *Collaborating for collaborative learning of text comprehension.* Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 2,* 117-175.

Palincsar, A. S., & Brown, A. L. (1985). *Reciprocal teaching of comprehension strategies: A natural history of one program for enhancing learning* (Tech. Rep. No. 334). Urbana-Champaign: University of Illinois, Center for the Study of Reading.

Palincsar, A. S., Stevens, D. D., & Gavelek, J. R. (1989). Collaborating with teachers in the interest of student collaboration. *International Journal of Educational Research, 13,* 41-53.

Paris, S. C., Cross, D. R., & Lipson, M. Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76,* 1239-1252.

Pearson, P. D., & Dole, J. A. (1987). Explicit comprehension instruction: A review of research and a new conceptualization of instruction. *Elementary School Journal, 88,* 151-165.

Raphael, T. E., & Pearson, P. D. (1985). Increasing student awareness of sources of information for answering questions. *American Educational Research Journal, 22,* 217-237.

Resnick, L. B. (1987). *Education and learning to think.* Washington, DC: National Academic Press.

Rich, R. Z. (1989). *The effects of training adult poor to use text comprehension strategies.* Unpublished doctoral dissertation, Teachers College, Columbia University, New York.

Risko, V. J., & Feldman, N. (1986). Teaching young remedial readers to generate questions as they read. *Reading Horizons, 27,* 54-64.

Rush, R. T., & Milburn, J. L. (1988, November). *The effects of reciprocal teaching on self-regulation of reading comprehension in a postsecondary technical school program*. Paper presented at the annual meeting of the National Reading Conference, Tucson, AZ.

Scardamalia, M., & Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions*. Hillsdale, NJ: Erlbaum.

Shortland-Jones, B. (1986). *The development and testing of an instructional strategy for improving reading comprehension based on schema and metacognitive theories*. Unpublished doctoral dissertation, University of Oregon, Eugene.

Taylor, B. M., & Frye, B. (1988). *Skill pretests: Replacing unnecessary skill activities with pleasure reading and comprehension strategy instruction*. Unpublished document, University of Minnesota, Minneapolis.

Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life*. New York: Cambridge University Press.

Williamson, R. A. (1989). *The effect of reciprocal teaching on student performance gains in third grade basal reading instruction*. Unpublished doctoral dissertation, Texas A&M University, College Station.

Author Note

This research was supported by the Bureau of Educational Research, College of Education, University of Illinois at Urbana-Champaign. Special thanks are extended to those colleagues who assisted us during the writing of this review: Patti Alexander, Philip Brady, Courtney Cazden, Saul Chapman, Lynn Corno, Margaret Dermody, Mark Gall, John Helfeldt, Fran Lehr, Anthony Manzo, Annemarie Palincsar, Mike Pressley, Rebecca Rich, Bill Rupley, Jo Shepard, Bob Slavin, and Donna Weisman.

Table 1

OVERALL RESULTS

Effect Size for All Studies[a]

	S	S/NS	NS	Median[b] Effect Size
All Studies	7	5	7	.57 (13)

[a]Each study was counted only once in this table.

[b]The number in parentheses after each effect size is the number of studies on which it is based.

Table 2

ADDITIONAL ANALYSES

Effect Size for All Studies[a]

	S	S/NS	NS	Median[b] Effect Size
All Studies by Instructional Approach				
ET/RT	4	5	1	.60 (8)
RTO	3	0	6	.34 (5)
All Studies by Type of Student				
Good-poor/all	5	4	3	.57 (9)
Below-average	2	1	4	.48 (4)
All Studies by Type of Test				
Standardized	4	0	9	.32 (9)
Experimenter Developed	8	1	2	.87 (9)

[a]Each study was counted only once in the analyses by type of instructional approach and type of student. In the section showing results by type of test, studies using both types of test were included in both categories and, therefore, counted twice.

[b]The number in parentheses after each effect size is the number of studies on which the effect size is based.

Table 3

Effect Sizes for Good-Poor/All Students[a]

	S	S/NS	NS	Median[b] Effect Size
All Studies				
ET/RT	3	4	1	.57 (7)
RTO	2	0	2	.59 (2)
Total	5	4	3	.57 (9)
Standardized Tests				
ET/RT	2	0	3	.32 (5)
RTO	1	0	1	.84 (1)
Total	3	0	4	.44 (6)
Experimenter-Developed Comprehension Measures				
ET/RT	5	1[c]	1	.93 (6)
RTO	1	0	1	.34 (1)
Total	6	1	2	.85 (7)

[a]Each study was counted only once in each of the appropriate tables. If a study had two dependent measures for a specific category, an average effect size was used.

[b]The number in parentheses after each effect size is the number of studies on which it was based.

[c]Grade 4 was significant; Grade 7 was not.

Table 4

Effect Sizes for Below-average Students[a]

	S	S/NS	NS	Median Effect Size[b]
All Studies				
ET/RT	1	1	0	.87 (1)
RTO	1	0	4	.14 (3)
Total	2	1	4	.48 (4)
Standardized Tests				
ET/RT	1	0	1	.36 (1)
RTO	0	0	4	-.21 (2)
Total	1	0	5	.14 (3)
Experimenter-Developed Comprehension Measures				
ET/RT	1	0	0	.87 (1)
RTO	1	0	0	1.44 (1)
Total	2	0	0	1.16 (2)

[a]Each study was counted only once in each of the appropriate tables. If a study had two dependent measures for a specific category, an average effect size was used.

[b]The number in parentheses after each effect size is the number of studies on which it was based.

Table 5

Effect Sizes for All Studies[a] by Type of Test

	S	S/NS	NS	Median[b] Effect Size
Standardized Tests				
ET/RT	3	0	4	.34 (6)
RTO	1	0	5	.14 (3)
Total	4	0	9	.32 (9)
Experimenter-Developed Comprehension Measures				
ET/RT	6	1	1	.87 (7)
RTO	2	0	1	.89 (2)
Total	8	1	2	.87 (9)

[a]Each study was counted only once in each of the appropriate tables. If a study had two dependent measures for a specific category, an average effect size was used.

[b]The number in parentheses after each effect size is the number of studies on which it was based.

Table 6

Grade Level of Students

Good-Poor/All			Below-average		
S	S/NS	NS	S	S/NS	NS
6-8	4 & 7	3	4-6	7th	3-5
4-5	5-6	3	adult		5
4 & 6	4	Voc. Col.			7-25 yr. olds
5	1-6				8-12 yr. olds
7					

Table 7

Median Number of Instructional Sessions

ET/RT			RTO		
S	S/NS	NS	S	S/NS	NS
20	17	20	14	25	24

Table 8

Instructional Group Size
Medians

Good-Poor/All			Below-average		
S	S/NS	NS	S	S/NS	NS
11.5	10	5	14	6	4

Table 9

Median Number of Cognitive Strategies Taught

Good-Poor/All			Below-average		
S	S/NS	NS	S	S/NS	NS
1	2	4	10	4	1
1	3	4			4
2	4	4			4
4	4				4
4					
2	3.5	4	10	4	4

Table 10

Person Providing Instruction

Experimenter			Teacher		
S	S/NS	NS	S	S/NS	NS
3	3	3	3	3	3

Table 11

Effect Size for Studies With Acceptable Control Groups

	S	S/NS	NS	Median[a] Effect Size
All studies with acceptable controls	6	2	6	.40 (8)

[a]The number in parentheses after each effect size is the number of studies on which it is based.

Table 12

Teacher Implementation

1. Palincsar (1987)

Two raters, who had conducted reciprocal teaching instruction, listened to an audiotape and rated the quality of instruction. "The raters noted that each of the teachers indicated fidelity to the reciprocal teaching procedure" (p. 18). The specific criteria used to judge the quality were not given in the study.

2. Palincsar & Brown (1984)

A random selection of audio tapes from five different sessions with each of the three groups allowed an independent evaluator to carry out procedural reliability checks. The evaluator "verified that the procedure was indeed implemented as described" (Palincsar, 1982, p. 74). The specific criteria used to judge the quality were not given in the study.

3. Manzo (1969)

Data were collected on the frequency and the type of questions asked by teachers and students in both the experimental and control groups. The six categories of questions were as follows:

(1) recognition
(2) recall
(3) translation
(4) conjecture
(5) explanation
(6) evaluation

Reciprocal teaching teachers were higher than control teachers in percentage of questions coded as translation, conjecture, and explanation, and lower than control teachers in questions coded as recognition and recall.

4. Taylor & Frye (1988)

Each week, one of the two researchers met individually with each teacher for 5 to 15 minutes to talk about the study, answer concerns, and answer questions. "Based on these discussions it was felt that the teachers were following procedures as requested." No specific rating items were given (p. 15).

5. Levin (1989)

There were three full-period observations, by the experimenter, of the control teachers. Four variables were observed:

(1) students leading the lessons
(2) students formulating questions about the material
(3) mention of ISL strategies
(4) basal or phonics book in use

For the control classes, no occurrences of variables 1, 2, or 3; variable 4 was consistently observed.

Experimental teachers were observed four times for 45 minute observations using an ISL Observational Checklist developed by Paris (1984). The checklist included 12 teaching behaviors:

(1) gains students' attention
(2) introduces lesson
(3) states goals

(4) reviews previous information
(5) informs students about (a) what it is, (b) how it works, and (c) when and when not to use
(6) encourages metaphorical reasoning
(7) persuades why strategy is useful
(8) models and demonstrates strategy
(9) allows immediate practice
(10) provides praise and corrective feedback
(11) fades support
(12) bridges to other reading/learning (p. 185)

The experimental teachers had difficulty with behaviors 6, 7, and 8. Although these three behaviors were infrequently observed at the beginning of the program, they were observed most of the time during the end of the program. Remedial work with the teachers was conducted throughout the program. No data were provided on the other nine behaviors found on the ISL checklist.

In addition to observations based on the above ISL checklist, the researcher made three observations of each class during reciprocal teaching sessions and recorded anecdotal field notes. These observations appear to have been focused more on the reciprocal teaching procedure than on the ILS behaviors. The notes revealed two types of teacher error.

Some teachers moved too quickly to take over when students were having problems formulating questions or working out strategies for getting classmates' cooperation (Teachers 1, 3A, 4, and 5). Others (Teachers 2 and 3A) did not interfere even when students made mistakes, but allowed students to mispronounce important words or to accept incorrect answers (p. 81).

Discussion and review of reciprocal teaching protocols and additional demonstrations of the procedure succeeded in helping most of the teachers overcome the problems mentioned above.

6. Shortland-Jones (1986)

The experimenter developed a checklist and observed teachers in four stages of a lesson:
(1) introduction/motivation
(2) introduction of the graphic organizer
(3) students generation of metacognitive strategies
(4) lesson conclusion

"It appears that both groups of teachers [for experimental and control groups] achieved a high rate of implementation of techniques in the introductory stage and in the graphic overview stage" (p. 71).

7. Williamson (1989)

All participating teachers (for both control and experimental groups) were observed by the researcher on three different occasions during regular reading instruction. A checklist containing criteria items related to each of the four reciprocal teaching activities was given to experimental teachers prior to implementation and was used during the observations. The actual checklist or a discussion of the criteria, however, was not provided. Individual conferencing was provided for any experimental teacher having difficulty adhering to the guidelines of reciprocal teaching.

The researcher reported a low degree of implementation overall.

> It was observed that not all teachers were equally effective in applying the reciprocal teaching strategy activities. Suggestions and additional coaching did not adequately remedy the situation. (p. 64)

The actual weaknesses found during implementation observations were not discussed in the study.

Table 13

Assessing Student Learning of the Strategies

1. Asking students to generate questions about a new passage
   - Dermody, 1988
   - Lysynchuk et al., 1990
   - Padron, 1985
   - Palincsar, 1987
   - Palincsar & Brown, 1984
   - Shortland-Jones, 1986
   - Taylor & Frye, 1988

2. Coding student questions during sessions
   - Manzo, 1969

3. Asking students to summarize a passage
   - Dermody, 1988
   - Lysynchuk et al., 1990
   - Palincsar, 1987
   - Palincsar & Brown, 1984
   - Shortland-Jones, 1986
   - Taylor & Frye, 1988

4. Asking students to make predictions
   - Dermody, 1988

Table 14

Prompts/Facilitators for Questioning and Summarizing Taught in Different Reciprocal Teaching Studies

1. Question-generation

   signal words (Brady, 1990; Lysynchuk et al., 1990; Palincsar, 1987; Palincsar & Brown, 1984; Taylor & Frye, 1988; Williamson, 1989)

   question types (Dermody, 1988)

   locate main idea (Lonberger, 1988; Rich, 1989)

   story grammar (Nolte & Singer, 1985)

2. Summarization

   Kintsch-van Dijk procedures (Brady, 1990; Palincsar, 1987)

   legs-and-table procedures (Dermody, 1988; Lonberger, 1988; Taylor & Frye, 1988)

APPENDIX A

STUDIES THAT USED RECIPROCAL TEACHING ONLY

Study	Length	Level	Type of Student	Dependent Measure	Sign.	Effect Size
Fischer Galbert, 1989	100 sessions[a]	3rd, 4th, and 5th	Chapter 1	Iowa TBS	NS[b]	-.21 (est.)
Helfeldt & Lalik, 1976	14 sessions	5th	average	Van Wagenen Anal. Reading Scales	S	.84
Jones, 1987	20 sessions	3rd	all students	Informal reading inventory	NS	.34[c]
Labercane & Battle, 1987	28 sessions	5th	learning disabled	Gates-MacGin.	NS	.36 (est.)
Manzo, 1969	30 sessions[d]	7-25 years old	remedial (summer tutorial)	Gates-MacGin.	NS	.36 (est.)
Padron, 1985	8 sessions	8-12 year olds	Hispanic bilingual	Stan. DRT (comp. subtest)	NS	-.55
Palincsar & Brown, 1984	20 sessions	7th	poor comp./ good decoders	comp. quest. (open ended)	S	1.0 + (est.)
Rich, 1989	6 sessions	adult	poor readers	Mult. choice comprehension	S	1.74[e]
				Recall (idea units)	S	1.10[e]
Rush & Milburn, 1988	9 sessions	voc. college	all	Cooperative English test	NS	no SD; 4 treatments

[a]number represents the total number of Chapter 1 classes.

[b]may favor control.

[c]average of two experimental groups.

[d]number of sessions for total reading instruction program.

[e]average of three experimental groups.

APPENDIX B

STUDIES THAT USED EXPLICIT TEACHING PLUS RECIPROCAL TEACHING

Study	Length	Level	Type of Student	Dependent Measure	Sign.	Effect Size
Signal Words Facilitator						
Brady, 1990	25 sessions	7th	below-average (native Alaskan)	Comp. test - open ended	S	.87
				Gates-MacGin.	NS	.36
Lysynchuk et al., 1990	13 sessions	4th and 7th	poor comp./ good decoders	Gates-MacGin./ Metropolitan comp. test	NS	.55
					NS/S	.68
Palincsar, 1987	25 sessions	middle school	poor comp./ good decoders	summarizing (levels)	S	.68 (est.)
				comp. ques. (open ended)	S	1.08 (est.)
Taylor & Frye, 1988	11 sessions	5th and 6th	average and above	Gates-MacGin.	NS	.07
				summarizing (levels)	S	.85
Williamson, 1989	50 sessions[a]	3rd	all students	IL State Assess.	NS	.32
Story Grammar Facilitator						
Nolte & Singer, 1985	10 sessions	4th and 5th	all students	mult. choice comp. test	S	1.0
Question Type Facilitator						
Dermody, 1988	24 sessions	4th	all students	Stan. DRT	NS	-.32
				summarizing (total prop.)	S	3.37
Main Idea Facilitator						
Lonberger, 1988	20 sessions	4th and 6th	all students	CR test (exper. developed)	S	1.24
Graphic Organizer						
Shortland-Jones, 1986	20 sessions	1st-6th	mixed (summer reading program)	Stan. DRT	S	.77
				summarizing	NS	-.02
Informed Strategies for Learning						
Levin, 1989	50 sessions	intermed.	learning disabled	Stan. Achiev. comp. subtest	S	no adj. means

[a]number represents total regular reading instructional periods.

APPENDIX C

The following is a script illustrating the explicit teaching of questioning prior to the reciprocal teaching dialogue. This script was developed by Annemarie Palincsar.

QUESTIONING
Intermediate Level

"Questions play an important part in our lives. Much of our class discussion is focused on answering questions. What are some other occasions, or situations when questions are important?"

Elicit responses from the students that might include the following:

- Reading assignments usually require us to answer questions at the end of a story.

- Tests usually require us to answer a series of questions.

- When we need more information about something, we need to ask questions.

"Let's practice asking some questions for situations when we might need more information. For example, suppose you want to see the movie, 'The Empire Strikes Back,' however, you don't know when the movie begins. You might call the theater and ask (What time does the movie, 'The Empire Strikes Back' begin?

"One of the activities we will be doing for the next couple of weeks is learning to ask good questions about material we have read. We will focus on asking questions about important information rather than about unimportant, trivial, or detailed information.

"There are several reasons why we will learn to ask questions while reading: (list underlined phrases on the chalkboard)

1. It is a way in which we can test ourselves to make sure we understand what we have read.

2. It is a good way to focus on important information in a passage.

3. With a little practice it is possible that we can become skilled enough at questioning that we can predict the kinds of questions we might be asked on a test. This would be very useful while studying.

"Let's begin by talking about the words that are used to ask questions. What are some of the words that we use to begin sentences that are questioning sentences?" (List responses on the chalkboard).

Who	What
When	Where
Why	How

"Let's practice by asking questions about the following sentences. At first you will be given the question word, however, later you will be asked to think of your question words. Look at the first sentence on your papers, 'The falcon is a female hunting bird.' Ask a question word about the information in this sentence that begins with the word 'what'."

1. The falcon is a female hunting bird.

   What is a falcon? /or/ What is the name of a female hunting bird? /or/ What does a falcon do? (Accept any appropriate responses)

"Ask a question word about the information in sentence #2. Begin your question with the word 'who'."

2. In medieval times, in Europe, only members of a royal family could own falcons.

   Who could own falcons in medieval times?

"Ask a question about the third sentence that begins with the word 'why'."

3. The falcon bathes in shallow streams to control bird lice that live in her feathers.

   Why does the falcon need to bathe?

4. A falcon prefers to hunt for its prey in open areas.

   Where do falcons hunt?

5. In the 1950s the falcon populations in North American and Central Europe dropped suddenly.

   When was there a decline in the falcon population?

6. The falcon hunts by swooping down on her prey and grabbing it with her sharp talons.

   How does the falcon catch her prey?

"For further practice, make up questions for each of the following sentences #7 through #11. This time, however, no question words are given." You may continue to complete these items as a whole-class discussion or give the students the opportunity to complete them independently by writing down their responses. Discuss their questions and accept any question that captures the main idea of the sentence and is posed clearly.

7. Although animals don't have language as we do, they do communicate with each other by signals of some kind.

8. Scientists study animal communication through experiments and observations.

9. Because snakes are totally deaf, it is the movement of the snake charmer that charms the snake, not the music the snake charmer plays.

10. Some ants give off a special alarm odor that warns nearby ants of danger.

11. The sounds made by bats, moths, and whales are too high for humans to hear.

"Now that you are successful making up questions, we will discuss selecting the most important information in the paragraph about which to ask a questions. Look at #12 on your papers. #12 is a short paragraph. After the paragraph there are three questions. One of these questions is better than the other two because it is about the most important information in the paragraph. Let's first read the paragraph."

12. Deaths from snakebite have been cut down in recent years by the use of antivenins--medicines that work against the snake poisons. There are now few deaths from snakebite in the United States and Canada.

"Let's read the three questions that follow this paragraph and try to decide which question asks about the most important information." (Read through all three choices).

___ a. Why do snakes bite people?

"This is not a good main idea question. Can anyone tell us why?" (This is not a good main idea question since the question is not answered in the paragraph. This question would make a good prediction, however, since it shows that you are thinking about what kinds of information might come next in the story.)

___ b. In what countries do few people die from snakebite?

"This is what we could call a detail question. While the answer is in the paragraph it is not about the most important information in the paragraph."

_X_ c. Why do fewer people die from snakebite these days?

"This question is the best because to answer it you must discuss the antivenin or medicine which is the main topic of the paragraph."

"Let's try another example. Read paragraph #13 and the three questions that follow it. Put a check mark next to the best question. Remember that the best question should be clear and should be about the most important information in the paragraph. Be ready to discuss your choice." Once students have selected the best response, discuss each question and why it is/is not the best choice.

13. Contrary to what some people believe, snakes do not sting with their tongues. Their tongues are used to sharpen their sense of smell. The snake picks up tiny particles of matter in the air with his tongue and puts them in two tiny holes at the bottom of his nostrils so that he can smell better.

___ a. How many holes does a snake have at the bottom of his nostrils? (detail question)

_X_ b. What does a snake use his tongue for?
(main idea question since paragraph discusses how a snake uses his tongue for smelling rather than for stinging)

___ c. Why do people use the expression, "he speaks with forked tongue?" (question not answered in the text)

Instruct students to select the best main idea questions for #14 to 16. Tell them to be ready to defend their answer by telling the reasons for their choice. Discuss student responses.

14. The smallest snake is just about the size of a worm. The largest snake has been known to reach thirty feet in length which is almost as long as two station wagons. There are many varieties of snakes and they come in many different lengths.

_X_ a. How long do snakes get? (all sentences discuss length of snakes)

___ b. How many station wagons could fit into 30 feet? (not even mentioned in paragraph)

___ c. Where would you find the longest snake? (not discussed in paragraph, a question like this would make a good prediction for the next portion of text)

15. Snakes are very flexible because their body is like a rubber hose with many bones. In fact, a snake's backbone can have as many as 300 vertebrae, almost ten times as many as a human's. Because of all these bones, a snake can twist its body in almost any direction.

_X_ a. Why can a snake move its body in so many ways? (entire paragraph is about flexibility and movement of snakes)

___ b. Do snakes ever need backrubs the way people do? (not mentioned)

___ c. How many vertebrae do snakes have? (detail question)

16. While very small snakes eat very small insects or worms, large snakes can eat small deer, leopards, and goats. All snakes, regardless of size, eat living animals or animal eggs. In fact, some snakes swallow each other.

___ a. What snake eats its neighbor snake? (not mentioned in paragraph)

_X_ b. What do snakes eat? (good main idea question since entire paragraph describes what snakes eat)

___ c. How is the diet of the small snake different from the diet of the snake? (a good question, however, this question is only about one sentence rather than the entire paragraph)

Instruct students to write a main idea question for #17 to #20. Discuss responses.

17. Camels have been helpful to people who live in deserts for thousands of years. They have carried people as well as their goods on their strangely shaped backs. They are able to cross deserts and mountains on trips that may take two months.

(Why or how are camels helpful to people?)

18. Scientists have studied the camel carefully to determine how it can live where other animals would die. They have found that the camel's body is especially well designed for its life in the hot, dry, sandy parts of the world. There are many characteristics of the camel that are useful to it including its feet, legs, eyelashes, and nostrils.

(Why is the camel able to live in the desert?)

19. There have been many prominent women in America's history who have done much good for mankind. One of these women was Alice Hamilton. Dr. Hamilton was very concerned about the health of industrial workers. Through her research and leadership she was responsible for many changes that improved working conditions for laborers.

(What did Dr. Alice Hamilton do to help people? or What was important about Dr. Alice Hamilton?)

20. Scientists have been asking themselves what energy is for hundreds of years, but no one has come up with a simple answer. About the only definition of energy that scientists can agree on is that energy is that something which enables people, machines, and objects to do work.

(What is energy?)

Source: Palincsar, 1987.